AMENDMENT(S) TO THE SPECIFICATION

Please replace the paragraph beginning at page 2, line 18, with the following rewritten paragraph:

Still another approach would be to low pass (or band pass) filter the high resolution data streams before attempting the cross-correlation. In this case, the cross-correlation function can be sampled at the lower resolution without worrying about the Nyquist rate: the low pass (or band pass) filter of the inputs into the cross-correlation function ensures that the Nyquist requirements are met. However, low pass (or band pass) filtering the input data so severely is likely to remove many of the distinctive identifying characteristics of the high resolution data streams, thus degrading the ability of the cross-correlation to produce accurate alignment. For example, if this approach is used with two audiovisual data streams, even if a "good" band is selected to pass, there are not many distinguishing features left in an audio signal that has been filtered down to a 15 Hz bandwidth (15Hz = ~~30Hz/2)~~ 30Hz/2, since sampling occurs at 30 Hz and Nyquist requires 2 samples/cycle).

Please replace the paragraph beginning at page 7, line 6, with the following rewritten paragraph:

In summary, if the audio stream is x[n], the global mean energy, A_{var}, is established according to the following equation:

$$A_{var} = MAX\left(\left(\frac{1}{T}\sum_{N=0}^{T-1} m_R\{x,N\} \right)^2 , \frac{1}{T}\sum_{N=0}^{T-1} v_R\{x,N\} \right) \quad (1)$$

where $m_R\{x,N\} = 1/R \sum_{n=0}^{R-1} x[n+RN]$ (2)

$$v_R\{x,N\} = \frac{1}{R}\sum_{n=0}^{R-1} x^2[n+RN] - m_R^2\{x,N\} \quad (3)$$

R = length of the short time window
RT = length of the long time window
T = constant relating the length of the long time window to the length of the short time window

The audio segments on which to cross-correlate are selected from the set of segments that satisfy the ~~follow~~ following conditions (the outer local mean and outer local variance estimates are taken using "N" as the sequence index):

$$m_S\{v_R(x,N),k\} > T_{level} A_{var} \quad (4)$$
$$v_S\{v_R(x,N),k\} > T_{var} A^2_{var} \quad (5)$$

where RS = length of the mid-length time window
S = constant relating length of the mid-length time window to length of the short time window
T_{level} = constant establishing threshold audio energy level for audio segment to be identified as distinctive
T_{var} = constant establishing threshold audio energy variance for audio segment to be identified as distinctive